WHAT IS CLAIMED IS:

1. In an information processing system, a method for speech processing, the method comprising:

receiving an utterance;

computing a score based on the utterance, including evaluating states of a model of a keyword; and

indicating based on the score that the utterance appears to contain the keyword;

wherein, in the computing step, the score is computed without requiring that a model, of speech other than the keyword, be evaluated only at states corresponding to the evaluated states of the model of the keyword.

2. The method of claim 1 wherein the computing step includes:

evaluating a state j of the model of the keyword for each timeslice t of multiple timeslices of the utterance;

evaluating a state k of a model, of speech other than the keyword, at the timeslice t, wherein the state k is chosen to maximize or minimize a value without requiring that the state k equal the state j.

3. The method of claim 2 wherein the computing steps includes computing a value based on the expression:

$$\frac{b_j^c(o_t)}{\max_{k=1}^{N} b_k^a(o_t)}$$

where $b_j(o_t)$ is the observation probability in the state j at frame t; c indicates the model of the keyword; a indicates the model of speech other than the keyword; and N is a number of states in the model of speech other than the keyword.

4. A system for speech processing, comprising:

a processor;

a memory;

a model of a keyword;

a model of words other than the keyword; and

logic that directs the processor to read an utterance; compute a score based on the utterance and on the model of the keyword and the model of words other than the keyword; and indicate that the utterance appears to include the keyword;

wherein the score is based on portions, of the model of words other than the keyword, that do not necessarily correspond to portions, of the model of the keyword, that were used to compute the score.

5. The system of claim 4 wherein the logic is configured to direct the processor to evaluate a state j of the model of the keyword for each timeslice t of multiple timeslices of the utterance and to evaluate a state k of the model of words other than the keyword at the timeslice t, wherein the state k is chosen to maximize or minimize a value without requiring that the state k correspond to the state j.

6. The system of claim 5 wherein the logic is configured to direct the processor to compute a value based on the expression:

$$\frac{b_j^c(o_t)}{\max_{k=1}^{N} b_k^a(o_t)}$$

where $b_j(o_t)$ is the observation probability in the state j at frame t; c indicates the model of the keyword; a indicates the model of speech other than the keyword; and N is a number of states in the model of speech other than the keyword.

7. In an information processing system, a method for speech processing comprising:

receiving an utterance;

for each of multiple keywords, computing a score based on the utterance

for each of multiple keywords, comparing the score to a threshold, wherein the

threshold for one of the multiple keywords need not be the same as the threshold for another of the multiple keywords; and

indicating based on result of the comparison that the utterance appears to contain the keyword.

8. The method of claim 7 wherein the threshold for the one keyword and for the other keyword are each set using training data based on Bayes risk and on the conditional probability distribution function of the keyword discriminative function for the respective keyword.

9. A speech processing system, comprising:

a processor;

a memory;

logic that directs the processor to:

read an utterance;

for each of multiple keywords, compute a score based on the utterance and compare the score to a threshold;

wherein the threshold for one of the multiple keywords need not be the same as the threshold for another of the multiple keywords; and

indicating based on result of the compare that the utterance appears to contain a keyword.

10. The system of claim 9 wherein the threshold for the one keyword and for the other keyword are each set using training data based on Bayes risk and on the conditional probability distribution function of the keyword discriminative function for the respective keyword.

11. In an information processing system, a method for processing speech of a language having a syllabic character set, comprising:

maintaining models of syllables of the language, wherein syllables corresponding to

some characters of the character set are modeled using at least three subword units;

receiving an utterance;

computing scores based on the utterance and the models; and

indicating the detected existence of a word in the utterance based on the scores.

12. The method of claim 11 wherein the language is Chinese.

13. The method of claim 11 wherein the language is Mandarin Chinese.

14. The method of claim 11 wherein the three subword models are hidden Markov models and comprise a context-dependent initial model.

15. A speech processing system for performing recognition on speech of a language having a syllabic character set, the system comprising:

a processor;

a memory;

models of syllables of the language, wherein syllables corresponding to some characters of the character set are modeled using at least three subword units; and

logic that directs the processor to:

receive an utterance;

computing scores based on the utterance and the models; and

detecting existence of a word in the utterance based on the scores.

16. The system of claim 15 wherein the language is Chinese.

17. The system of claim 15 wherein the language is Mandarin Chinese.

18. The system of claim 15 wherein the three subword models are hidden Markov models and comprise a context-dependent initial model.